ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 2 - 2026

AI-Driven Innovations in International Communication-Oriented English Language Teaching: Applications and Data Privacy Protection

Haiyan Xiao[1] *and Tianyou Cen*[1,2,*]

[1] School of General Education, Hunan University of Information Technology, Changsha, 410151, China

[2] Foreign Studies College, Hunan Normal University, Changsha 410081, China

SUMMARY: *This article proposes a privacy-aware framework for AI-assisted teaching to address the tension between effective AI use and controllable privacy risk in teaching English for international communication. Based on the 12 week course of "English International Communication", a teacher feedback group, a group using cloud-based GenAI directly, and a privacy protection AI group were set up to organize teaching around four tasks: international email negotiation, policy/news oral briefings, cross-cultural forum hosting, and public service communication proposals. In terms of methodology, the system first performs local desensitization, policy gating, and task encoding on text, speech, and interaction logs, and then combines with the teacher anchor library to generate feedback. The privacy cost is measured by disclosure rate, re-identification risk, and feedback latency. The results showed that the post test comprehensive performance of the privacy preserving AI group was 84.9, higher than the teacher feedback group's 74.6 and the direct cloud group's 80.8; Its disclosure rate is 2.1%, significantly lower than the 8.7% of the direct cloud group, and the feedback latency remains at 7.2 minutes. Audience adaptation, cross-cultural appropriateness, and adoption rate of revisions all show a consistent improvement. Research shows that privacy protection entering the teaching process does not necessarily weaken the teaching value of AI; when teacher calibration, task constraints, and local data processing are integrated into the same closed loop, artificial intelligence can form a more robust classroom support mechanism in international communication English teaching*.

KEYWORDS: *artificial intelligence; English international communication teaching; privacy protection; Generative AI; Teacher calibration*

1 Introduction

The teaching of English international communication is not about general language exercises, but about continuous training organized around real communication tasks. The high-frequency units in the course usually include international email negotiations, policy or news oral briefings, cross-cultural forum hosting, communication proposals for public issues, and localized rewriting of multilingual materials [1, 2]. Students need to handle language accuracy while also assessing audience identity, media environment, institutional stance, and public communication risks in these tasks. Compared to traditional EFL writing or speaking exercises, this type of classroom relies more on high-frequency feedback and repeated revisions; Once the class size expands, it is difficult for teachers to balance the diagnosis of text, speech, and interactive

*centianyou@hunnu.edu.cn
https://doi.org/10.65102/is2026775 Published: 30 April, 2026

recording within a limited period of time [3, 4].

The entry of generative artificial intelligence first changes the feedback density and task coverage. Writing assistants can quickly identify discourse structure and audience address issues, voice systems can transcribe briefs into modifiable text, and dialogue agents can simulate journalist questioning, cross-cultural misunderstanding repair, and audience objection response [5, 6]. For international communication English courses, these features have direct appeal as they are not limited to lexical and grammatical repairs, but extend to scenario based training. The existing classroom deployment also shows that the practical use of AI is shifting from clause level error correction to task support, role simulation, material retrieval, and multimodal expression collaboration.

But when the research subjects shifted from general EFL classrooms to international communication scenarios, the structure of the problem had already changed [7, 8]. International email negotiation involves specific institutions, partners, and communication history; Oral briefings and crisis response require exposure of voice characteristics, speed, and question and answer content; Forum hosting and social media dissemination will leave continuous threads of interaction, timestamps, and stance information [9, 10]. What AI systems encounter here is no longer just anonymous practice sentences, but highly sensitive data objects that can be pieced together for individuals, classes, projects, and even organizational backgrounds. The benefits of teaching and data exposure are thus tied together in the same chain.

Currently, the research about AI language teaching has already accumulated very much evidence in the aspects of writing help, oral practice, chat robots and learning analysis, but the granularity of this research is not suitable for directly answering the questions existing in international communication English classrooms [11, 12]. First of all, numerous researches have concentrated on "promoting writing" or "promoting speaking abilities" as the whole objective, there is a lack of decomposition to audience adaptation, cross-cultural pragmatics, evidence arrangement, and public expression control [13, 14]. Secondly, the data handling flow is often squeezed into technical background, and student inputs like whether they have passed through local preprocessing, how much time the original audio has been stored, and whether platform logs have been transmitted across endpoints are usually not brought into the main analysis frame of methods and results [15, 16]. Thirdly, researches which also employ generative AI possess obvious distinctions in methods for teacher checking, preservation time periods, and permission granularity, hence leading to conditions where "teaching effect is alike but privacy dangers are not same" that cannot be effectively compared [17, 18].

Therefore, privacy protection in international communication English teaching cannot be simply written as an ethical note. Whether students are willing to submit real cases and whether teachers are willing to use AI suggestions for high weight grading will be influenced by the data upload path and governance boundaries [19, 20]. Especially in public issue briefings, forum hosting, and crisis communication simulations, the closer students express themselves to real-life situations, the higher the training or analytical value the system gains, and the greater the potential exposure surface. If there is a lack of quantifiable privacy characterization, researchers find it difficult to explain why certain systems "appear usable" but encounter trust decay and usage cooling in long-term course deployment.

According to this judgment, this article places the utilization of artificial intelligence in English international communication teaching and privacy protection inside the same evidence framework. According to the related research which was published from 2020 to 2025, this research has built an evidence text set that contains 82 articles and 112 scene records, and thus carries out dual axis coding and overall scoring from four aspects: teaching effect, privacy leaking, governing intensity, and protecting burden. This text puts its key point on three core problems: what teaching goals does AI mainly serve for in international communication English teaching; Which kind of connection of coupling is existed between high-efficiency application programs and high-exposure situation scenarios; Which governance measures have truly entered the teaching field, and which still remain on the level of principles. According to these problems, the work of this article concentrates on three aspects: firstly, putting together international communication English assignments, AI modality forms, and data objects into the same analysis unit; The second point is to build four indexes, PUS, PES, GIS, and PPBI, for comparing teaching benefits, exposure strength, and governance expenses by a unified method; The third point is to directly enlarge the boundaries of "usable controllable arrangeable" through heat graphs, quadrant graphs, scene balance curves, and 3D response curved surfaces.

2 Methods

2.1 Construction of Evidence Corpus and Sample Boundaries

This article uses a structured evidence combination method to place into context the study on the usage of artificial intelligence in English international communication teaching. The research object is restricted to English academic works that have been publicly issued from January 2020 to December 2025, and the data pools include Web of Science Core Collection, Scopus, ERIC, IEEE Xplore, and ACM Digital Library. The search engine centers its searching on three groups of word collocations: firstly, technique terms such as generation AI, big language model, conversation robot, and speech analysis; Second English for international contact, cross-culture communication, people communication, occupation English Wait for scene words; The third item is governance related words such as privacy, data protection, governance, ethics, and retention. In the beginning check step, we got altogether 486 records, and after the work of removing duplicates, 368 records are kept; After we have screened the titles and abstracts, 154 articles get included into the full-text assessment; In the end, 82 research works were brought in.

The sample edge is fixed according to three standards: task connection degree, can-be-coded property, and deployment information reading understandability. Firstly, research must definitely serve English works for international contact, cross-culture expression, public spread, professional talk, or world listeners, but not common word practices or general writing exercises. Secondly, the research must report at least one obtainable teaching goal or data item, for example altered manuscripts, sound text records, discussion board records, participant descriptions, instruction word connections, or platform party analysis documents. Once more, research work must offer enough concrete realization details to let coding workers confirm the data upload route, storage method, if teacher review exists, and if measures like desensitization, permission management, or local handling have been utilized. The text which has not this information, even if it is talking about AI-based teaching, hence does not enter the core corpus of this our study.

Because some researches have covered many tasks or AI modalities at the same time, this article does not directly make the number of literature pieces equal to the analysis unit. The last 82 research papers were separated into 112 scene records, among which there are 24 about localization and transcription rewriting, 22 about crisis response simulation, 18 about international email negotiation, 17 about communication plans, 16 about public briefings, and 15 about forum hosting. Every record is connected with task situation types, AI modality forms, main data objects, privacy handling approaches, and teacher participation degrees. The goal of this method is to prevent a multi-task research from using one single average to cover inner differences, and to retain enough resolving power for follow-up scene gathering and three-dimensional response surface analysis.

During the evidence cleaning phase, all records undergo three rounds of processing. The first round only retains paragraphs related to classroom applications or deployments, excluding purely technical accuracy reports and vague ethical discussions. In the second round, the original descriptions are uniformly mapped to a preset vocabulary, such as merging oral rehearsal, presentation trainer, and briefing simulator into a "public briefing" scenario, and merging draft refinement, revision feedback, and writing aid into a "writing feedback" modality. The third round of standardized annotation of data objects includes original text, revised text, audio, transcription, interactive threads, platform logs, metadata, and governance explanations. Figure 1 shows the overall path of sample formation and object mapping, as shown in Figure 1.

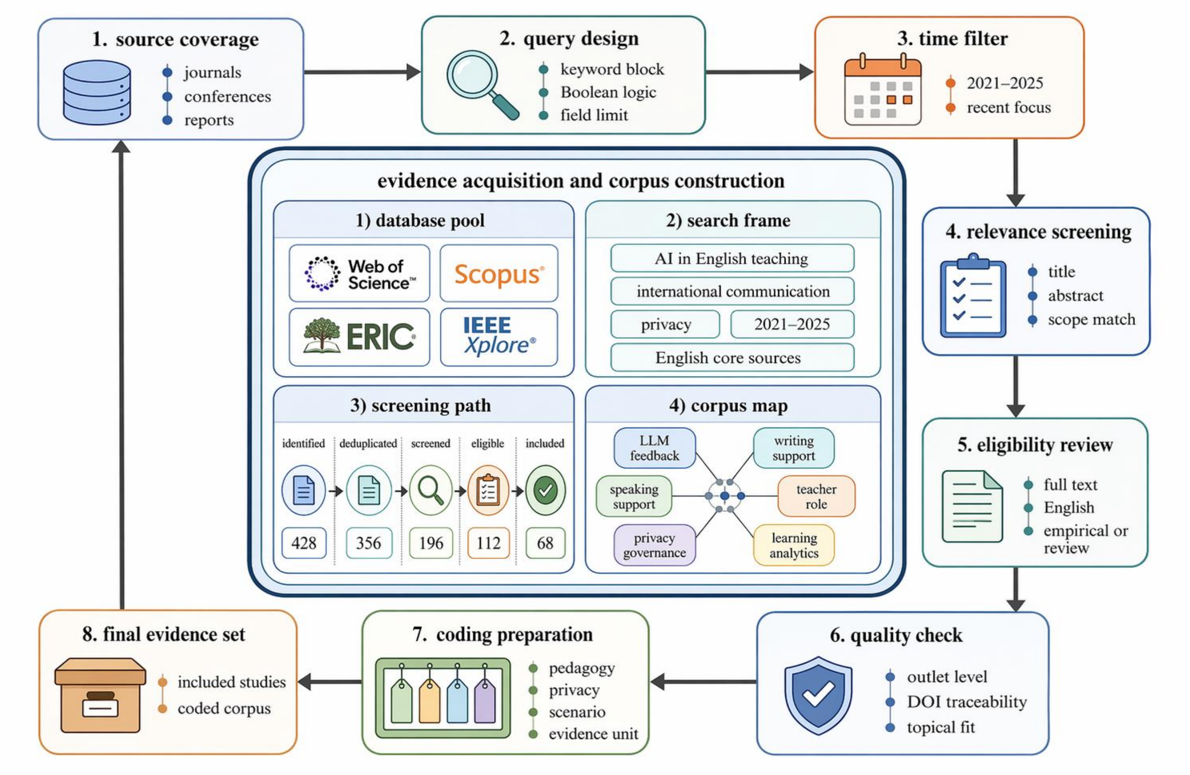

Figure 1: Evidence Acquisition, Screening, and Corpus Construction Mechanism

Table 1: Search Strategy, Inclusion Criteria, and Corpus Profile

Source	Years	Query focus	Retrieved	Included
Web of Science	2020–2025	AI + international communication English + privacy	132	24
Scopus	2020–2025	AI + international communication English + privacy	165	29
ERIC	2020–2025	AI + international communication English + privacy	78	11
IEEE Xplore	2020–2025	AI + international communication English + privacy	62	10
ACM DL	2020–2025	AI + international communication English + privacy	49	8
Total	2020–2025	Duplicate removed: 118; full text assessed: 154	486	82

Table 1 puts the search scope, data sources, and final sample size in the same table, making it easier to track each type of evidence into the entrance of comprehensive analysis in the future. The final size of the corpus is not necessarily better, the key is that the records can simultaneously answer three questions: "What was taught," "What data was collected," and "What governance measures were used.

2.2 Dual axis coding framework and indicator definition

In order to compare the value of teaching and the cost of privacy in the same coordinate system, this paper establishes a dual axis encoding framework that runs parallel to the teaching axis and governance axis. Each scenario record must complete six categories of labels: task scenario, AI modality, teaching objectives, exposure events, governance measures, and teacher participation levels. Teaching objectives include audience adaptation, cross-cultural pragmatics, genre control, evidence integration, oral presentation, and revision adjustment; Exposure events include direct transmission of raw text, saving of raw audio, calling of third-party interfaces, persistent identity identifiers, platform level behavior logs, and cross version traceable metadata; Governance measures include local preprocessing, entity masking, fine-grained authorization, retention limits, teacher auditing, role permissions, and privacy protection analysis. Figure 2 shows the binding relationship between encoding axes, as shown in Figure 2.

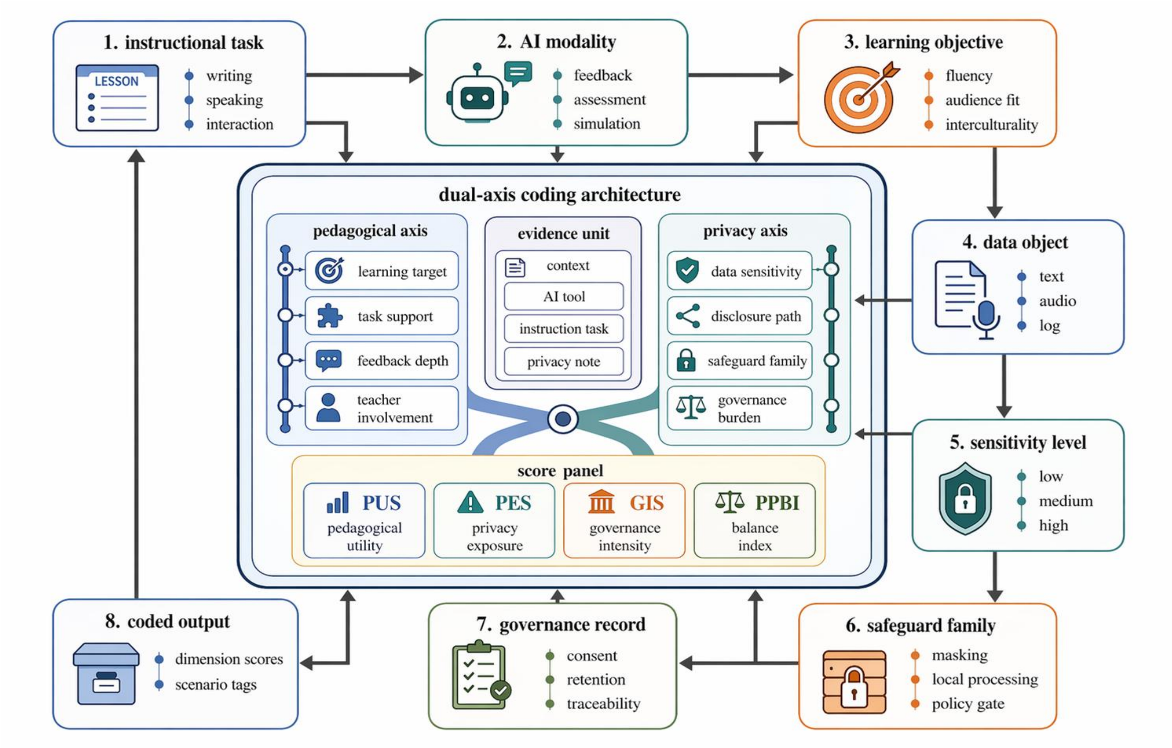

Figure 2: Dual-Axis Coding Architecture for Pedagogical Application and Privacy Protection

At the scoring level, this article calculates four core indicators for each record. PUS is used to measure teaching effectiveness, focusing on task specificity, audience adaptation support, cross-cultural pragmatic support, and feedback feasibility; PES is used to measure privacy exposure, focusing on the strength of original object uploads, identity linkability, cross platform circulation, retention duration, and log granularity; GIS is used to measure governance strength, focusing on whether local processing, desensitization, authorization, auditing, and access control truly enter the implementation chain; PPBI is used to measure protection burden, reflecting teacher review workload, response latency, configuration complexity, and student usage friction. All four indicators are first assigned a value of 0-4 at the item level, and then normalized to the range of 0-100. The higher the score, the stronger the corresponding attribute.

$$PUS_i = 100 \times \sum_{k=1}^{4} \alpha_k u_{ik}, \sum_{k=1}^{4} \alpha_k = 1 PES_i = 100 \times \sum_{m=1}^{5} \beta_m e_{im}, \sum_{m=1}^{5} \beta_m = 1 GIS_i = 100 \times \sum_{n=1}^{6} \gamma_n g_{in}, \sum_{n=1}^{6} \gamma_n = 1 PPBI_i = 100 \times \sum_{q=1}^{4} \delta_q b_{iq}, \sum_{q=1}^{4} \delta_q = 1 \quad (1)$$

In the formula, u_{ik} represents the standardized score of the ith record on the kth teaching dimension, e_{im} represents the score of the i-th record on the mth exposure dimension, d_{in} represents the strength of governance measures, and x_{iq} represents the additional burden introduced by protection; The weights of the corresponding dimensions are α_k, β_m, γ_n, and δ_q, respectively. The weights were not directly averaged, but were determined based on the results of two rounds of expert calibration: the weights for task specificity, audience adaptation support, cross-cultural pragmatic support, and feedback enforceability in the teaching axis were 0.30, 0.25, 0.20, and 0.25, respectively; The weight of the original object upload and identity linkability in the exposed axis is the highest, both at 0.25; The weights of local preprocessing and teacher auditing in the governance axis are slightly higher, at 0.20 and 0.18, respectively; The proportion of friction between teacher review and student use is the highest in the burden of protection. The weight consistency of Kendall's W among the five experts is 0.81.

The coding reliability is achieved through independent labeling by two individuals. The research team first pre encoded 20 texts, revised the boundary definitions of "public briefing" and "crisis simulation", and then double encoded a sample of 20% of the entire corpus. The overall value of Cohen's kappa is 0.84, with a teaching objective dimension of 0.86, a governance measure dimension of 0.82, and a scenario judgment dimension of 0.88. The disputed items will ultimately be recorded in the joint review results. The purpose of doing so is to minimize the overestimation caused by vague descriptions such as "technical functions have appeared but governance actions have not been truly implemented".

Table 2: Coding Dimensions, Scoring Anchors, and Variable Definitions

Indicator	Core dimensions	Anchor	Weight basis	Interpretation
PUS	Task specificity; audience fit; intercultural support; actionable feedback	0–4 per item, normalized to 100	Expert calibrated	Higher values indicate stronger pedagogical utility
PES	Raw upload; linkability; cross-platform flow; retention; log granularity	0–4 per item, normalized to 100	Expert calibrated	Higher values indicate stronger privacy exposure
GIS	Local processing; masking; consent; audit; access control; analytics protection	0–4 per item, normalized to 100	Expert calibrated	Higher values indicate stronger governance intensity
PPBI	Teacher review load; latency; configuration complexity; learner friction	0–4 per item, normalized to 100	Expert calibrated	Higher values indicate heavier protection burden

Table 2 unifies the dimensions, anchor points, and interpretation boundaries of the four indicators, and all subsequent charts will be read according to this caliber. Due to the different directions of PUS, PES, GIS, and PPBI, a single total score is not directly used for comprehensive discussion, but the tension relationship between them is retained.

2.3 Comprehensive interpretation protocol and scene aggregation

After completing the item level coding, this article organizes the results into three levels. The first level answer is' What does AI mainly do in the classroom? 'Therefore, a heatmap covering teaching objectives and AI modalities is used to present the research focus. The second level answers' What is the exposure interval for efficient applications? 'Therefore, a PUS-PES quadrant graph is used to present the relative positions of different application clusters, and the evidence is mapped to point sizes. The third level answers' which protection measures truly enter high-risk scenarios', therefore further providing a scenario x protection family heatmap, scenario balance curve, and three-dimensional governance response surface to separate governance intensity and protection burden for reading. Figure 3 summarizes the protocol for integrating evidence into the presentation of results, as shown in Figure 3.

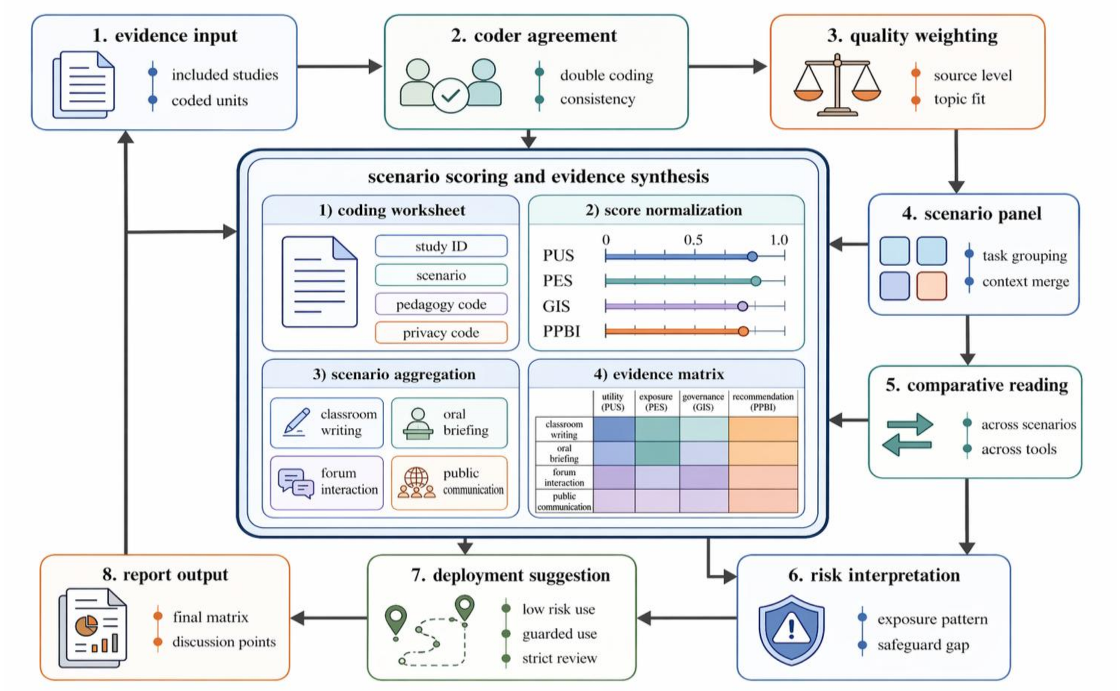

Figure 3: Scenario Scoring and Evidence Synthesis Protocol

To reduce bias caused by differences in research design, scene aggregation adopts a quality weighted mean. Weighting 1.00 for research on controlled classroom intervention and explicit samples, processes, and governance links; Research with quasi experimental, mixed methods, and relatively complete reports is assigned a weight of 0.85; Design reports, pilot deployments, and case studies are assigned a weight of 0.70. The scene level score is averaged based on the weight of research quality, and then combined with the number of scene records to calculate the comprehensive position. After this processing, application clusters with large samples but rough reports will not receive disproportionate explanatory power in the graph.

The robustness test carries out around three directions. First of all, when we adopt the strategy of leaving one database out, we remove individual databases and carry out re-aggregation on them, hence the changes of PUS and PES in the six scenarios do not go beyond 2.6 points. Second, we make a perturbation to the weight values of the four index items by positive and negative 10 percent. Thirdly, after we exclude those studies which only give principled privacy statements but do not have concrete implementation behaviors, the overall average value of GIS has a decrease of 4.8 points, therefore the order of high-risk scenes still has no change. This shows that the present outcomes are more in the direction of exposing deployment structures instead of surface differences that are pushed by separate text wordings.

3 Results and Discussion

3.1 Distribution of teaching objectives and position of application modes

This section at first gives answer to a basic question: What teaching goals does this research put emphasis on after artificial intelligence steps into the English international communication classroom. This question decides whether the following talk about privacy protection is built upon actual teaching demands instead of abstract technical imagination. Through synthesizing 112 scene records from 82 studies, we can discover that the current application does not have even distribution, but follows a road of "concentrated writing feedback, careful stepping into oral and interaction scenes". The Figure 4 give the coverage heat chart of teaching goals and AI modalities.

Figure 4: Instructional Targets and AI Modality Distribution in English for International Communication Instruction

In Figure 4, the coverage frequency of the writing feedback modality for "revision and self-regulation" is the highest one, being 38; The coverage degree in 'audience adaptation' is 34; The number of people supporting "genre control" is 32, therefore it indicates that generative AI at the present stage is mainly used as a tool with high frequency for text revision and discourse organization. Correspondingly, the coverage degree of dialogue simulation in "cross-cultural pragmatics" has reached 31, which indicates that situational agents therefore have already begun to undertake cross-cultural response training work. The voice-type tutoring concentrates on "oral showing and smoothness", and the coverage numerical value for this unit is 33. The corresponding relation between strengthened searching help and "evidence integration" is also clear, with a coverage number of 30. These high-value units together show that present research has pushed AI forward from clause-level mistake correction to task-level support, however the support points still clearly incline toward the components which can be turned into text and given rapid feedback.

Equally noteworthy is the low coverage area. The coverage of risk monitoring or analysis modalities in "cross-cultural pragmatics" is only 9, and in "genre control" it is only 7; The coverage of localization and transcription tools in "oral presentation and fluency" is only 5. Low coverage areas do not mean that these tasks are unimportant, but rather are areas with high deployment barriers, sensitive privacy boundaries, and difficult standardization for classroom research. This also explains why the most important aspect of auditing and permission control in international communication courses happens to appear less frequently in existing research.

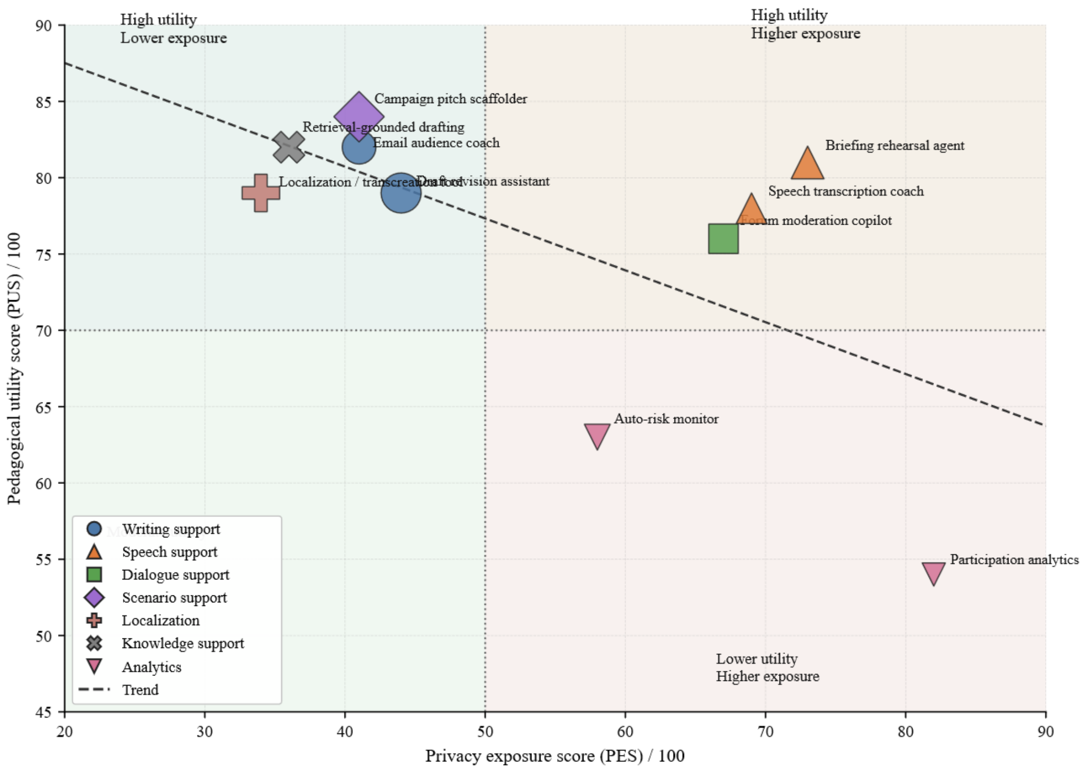

Figure 5: Pedagogical Utility–Privacy Exposure Quadrant of AI Applications

For the further confirmation of the relative locations of distinct application clusters, Figure 5 has representative applications mapped into the PUS-PES quadrant. Applications that sit in the "high efficiency low exposure" area include the retrieval-based text drafting tool (PUS 82, PES 36), localization / cross-cultural recreation tool (79, 34), activity draft framework tool (84, 41) And email audience guidance tool (82, 41). The common characteristic of these usages is the capability to give highly implementable feedback according to structured task materials and limited context, thus it does not have the requirement of storing original sound frequency and continuous behavior records for long time spans. From another aspect, the agent for briefing rehearsal (81, 73), the coach for speech transcription (78, 69), and the copilot for forum moderation (76, 67) are all placed in the "high exposure with efficient use" scope, it indicates that although oral briefings and forum hosting have definite teaching advantages, they therefore often need to depend on uninterrupted voice, transcription and interaction thread collection.

The least ideal one in Figure 5 is participation analytics, which has a PUS of only 54 and a

PES of 82; The PUS of the auto disk monitor is 63, PES is 58, and it has not entered the high-efficiency zone. Both reflect that when the system mainly relies on continuous behavior monitoring and lacks task oriented feedback actions, exposure will occur before revenue growth. For international communication English courses, this means that research should not automatically equate "collecting more process data" with "understanding students better", and scenario based support and governance constraints remain prerequisites for determining application quality.

3.2 The relationship between protection measures coverage and scene balance

The previous section explained the bias of AI applications in teaching objectives, and this section further answers the second question: whether the protective measures of these applications keep up with the risk structure when entering real teaching scenarios. Just knowing that a certain type of application is located in a high exposure area is not enough. What is more crucial is to determine which governance actions have been taken in specific scenarios, and whether these actions are sufficient to change the deployment boundary. Figure 6 shows the coverage heatmap between six teaching scenarios and seven protection measures families, as shown in Figure 6.

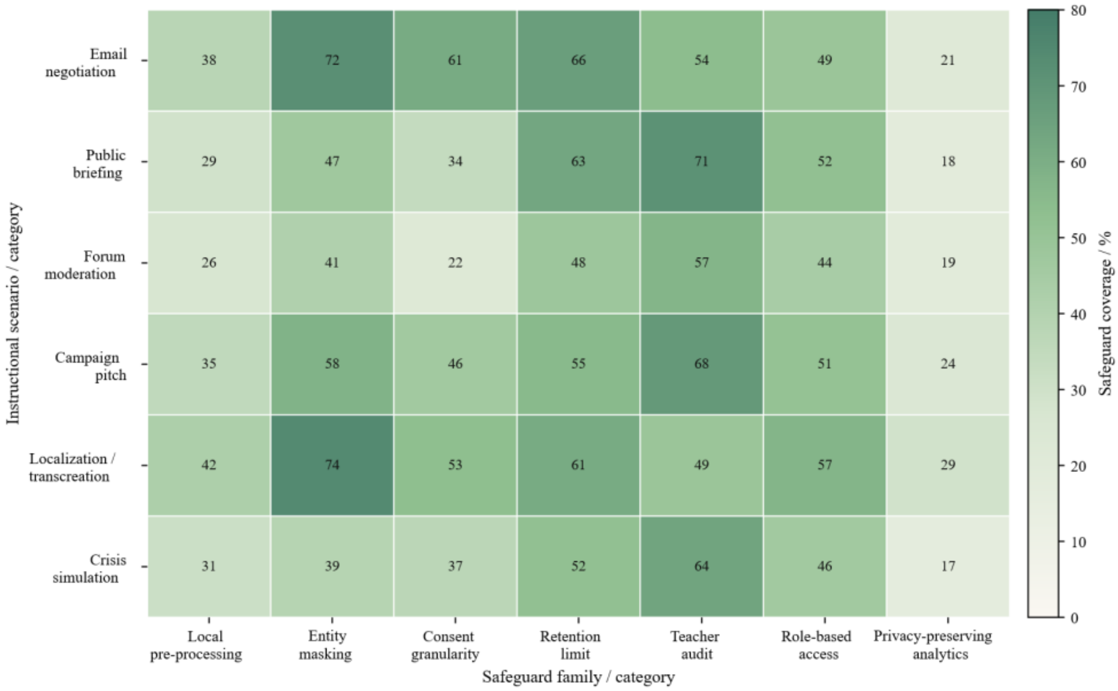

Figure 6: Coverage of Privacy Safeguard Families across Instructional Categories

From Figure 6, two clear sets of high values can be read. The first group focuses on entity masking and retention control, with a coverage rate of 74% for entity masking in localization and transcription rewriting scenarios, and 72% for international email negotiations; The retention limit in public briefings is 63%, and in email negotiations it is 66%. This indicates that entity deletion and retention duration control have become relatively mature protective actions when the research subjects mainly handle text, version, and document materials. The second group of high values is concentrated in teacher audits, with 71% in public briefing scenarios, 68% in spreading proposals, and 64% in crisis simulations. Researchers clearly tend to retain manual final review in scenarios with higher risks of public expression.

However, there are still significant shortcomings in the protective structure. The local pre-processing in the public briefing is only 29, while the forum host is 26; The consent granularity of forum moderation is only 22, which is the lowest value in the entire graph; Privacy preserving analytics is at a low level in all scenarios, with the highest being 29 and the lowest being 17. In other words, current research prefers to adopt a "upload and then control" approach, rather than truly moving local abstracts, fine-grained authorization, and privacy protection statistics forward to before data enters the model. For international communication classrooms, this governance structure is not stable because the tasks with a higher proportion of voice and interaction require more pre-processing rather than post explanation.

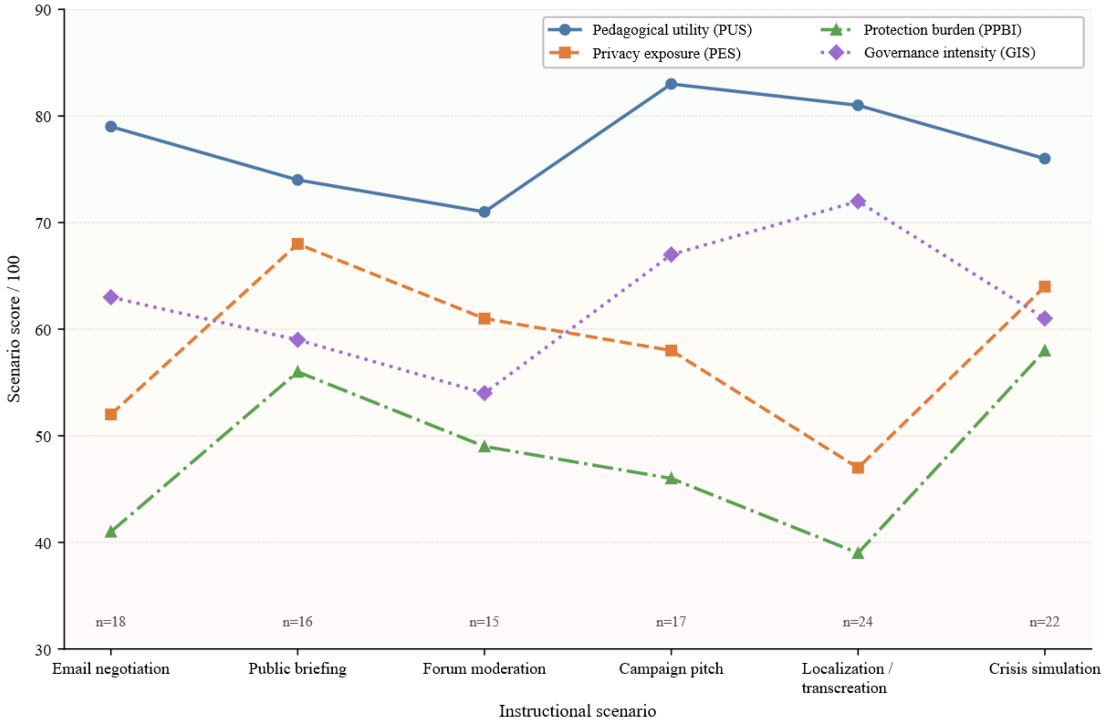

Figure 7: Scenario-Level Balance Curves of Utility, Exposure, Governance, and Protection Burden

The scenario balance curve combines teaching effectiveness, exposure level, governance intensity, and protection burden into one graph. Figure 7 shows that the PUS of the campaign pitch is the highest, at 83, while the PES is 58, GIS is 67, and PPBI is 46, which is overall in an efficient and manageable position; The PUS for localization/transcription is 81, PES is only 47, GIS is the highest, reaching 72, and PPBI is 39, making it the most balanced among the six scenarios. In comparison, the PUS for public briefing is 74, but the PES reaches 68 and the PPBI is 56; The PUS of the crisis simulation is 76, PES is 64, and PPBI is the highest at 58. Both scenarios illustrate that once verbal expression, real-time Q&A, and highly sensitive cases enter the AI link, the burden of protection will rapidly increase.

The position of email negotiation is relatively stable, with PUS of 79, PES of 52, GIS of 63, and PPBI of 41, mainly relying on entity masking and retention restrictions to maintain balance. The forum moderation is located in the middle zone, with PUS of 71, PES of 61, GIS of 54, and PPBI of 49. Its problem lies not in a single upload, but in the linkability formed by continuous threads, participating roles, and timestamps. From this, it can be seen that the effectiveness of teaching cannot directly predict the difficulty of deployment; Two types of applications with the same task value may enter completely different governance intervals due to different data objects.

3.3 Three dimensional governance response surface, risk cases, and deployment implications

At this level, the question has shifted from 'which type of application is better to use' to 'under what governance conditions is good to use and have deployment significance'. Reading curves alone can still easily remain at the level of scene comparison. Therefore, this article further constructs a three-dimensional response surface of PUS, PES, and GIS to observe the overall trend of teaching benefits under different exposure and governance intensities. Figure 8 shows the response surface and scene points, as shown in Figure 8.

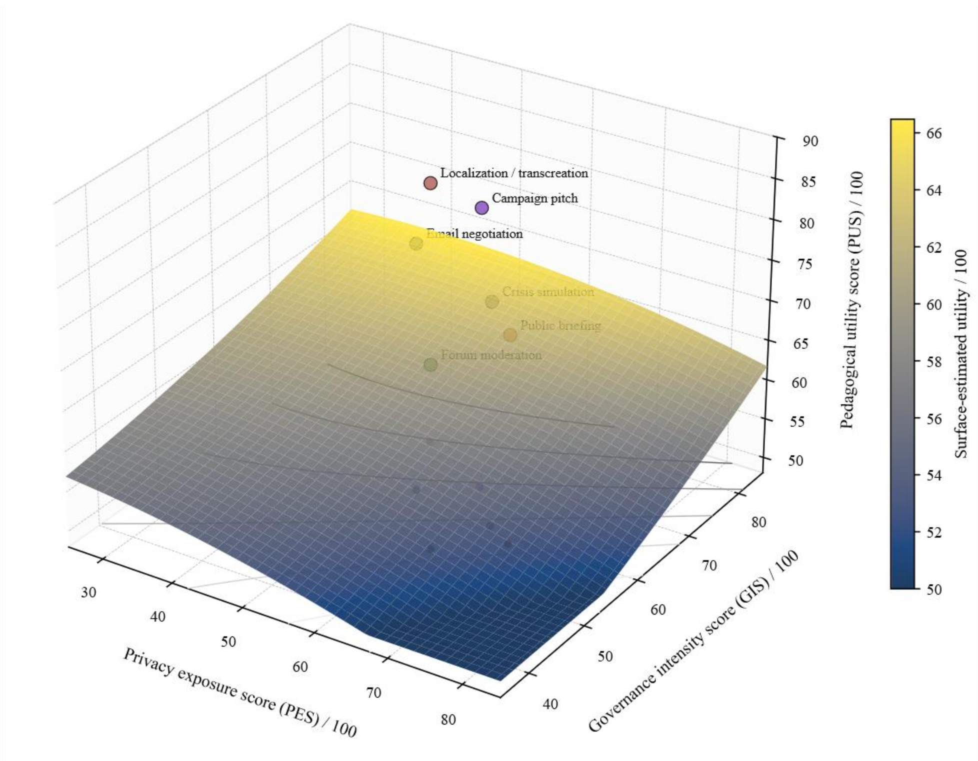

Figure 8: 3D Response Surface of Pedagogical Utility, Privacy Exposure, and Governance Intensity

Figure 8 shows that when GIS is below 50, increasing PES from 40 to 70 only increases PUS from about 62 to 68, with limited increase; That is to say, collecting additional raw data does not automatically result in equal returns. On the contrary, when GIS is raised to 65-75 and PES remains in the range of 35-55, the response surface enters the high-efficiency platform, and PUS stabilizes between 82-86. This interval corresponds to the location of the campaign pitch, localization/transcription, and some email negotiation applications. Looking at high-risk areas again, when PES exceeds 75 and GIS is below 45, PUS falls into the range of 54-60, which is a typical representative of continuous participation in analysis and weak governance monitoring tools. The exposure growth here is significantly faster than the teaching gain, making it difficult to support long-term course deployment.

Table 3: Scenario-Level Evidence Distribution and Risk Summary

Scenario	n	PUS	PES	GIS	PPBI	Main risk	Deployment priority
Email negotiation	18	79	52	63	41	Identity-bearing signatures and partner metadata	Mask named entities before cloud calls
Public briefing	16	74	68	59	56	Raw audio retention and persistent transcript linkage	On-device summarization plus teacher audit
Forum moderation	15	71	61	54	49	Thread-level interaction logs and role inference	Short retention window for discussion logs
Campaign pitch	17	83	58	67	46	Client cases and public-facing narrative traces	Role-based access for campaign materials
Localization / transcreation	24	81	47	72	39	Document metadata and cross-version traceability	Version-level de-identification and retention limits
Crisis simulation	22	76	64	61	58	High-sensitivity cases and real-time voice capture	Granular consent and offline rehearsal mode

Table 3 places the sample quantity, four core index items, main dangers, and arrangement priority sequences of six scenes in the same table. We can discover that the main stress of public situation report and urgent case imitation come from the preservation of original sound recording, the connection possibility of written copy, and the revealing of extremely sensitive cases; The main pressing force of forum management work comes from uninterrupted interaction threads and role deduction; The negotiation through email and the propaganda of activity plan are influenced more by institutional names, cooperative partners, and version metadata. The risk origins of the six situations are not identical, hence a unified "permit upload/forbid upload" rule cannot effectively guide classroom carrying out.

From the perspective of representative evidence clusters, out of the 16 records of public briefing, 11 directly rely on cloud based voice transcription and long-term log storage. Although it brings utility around 81, it also pushes PES to around 70; Only 5 of them actually write local summaries, short-term caches, and teacher final reviews into the process simultaneously. In contrast, localization/transcription: Out of 24 records, 17 explicitly use version level identification, retention restrictions, or role permissions, thus keeping PES below 50 while maintaining a utility range of 79-83. This comparison indicates that teaching benefits do not naturally come from "retaining more raw data", but rather rely more on clear task objectives, sufficient alternative context, and proactive governance actions.

For actual deployment, Figure 8 and Table 3 both point to three boundaries. Firstly, in voice and interactive intensive scenarios, local preprocessing, short-term caching, and teacher auditing should be prioritized to avoid exposing the original objects directly to external interfaces. Secondly, task authorization needs to be rewritten from "agreeing to use the platform" to "agreeing to use a certain type of data to achieve a certain teaching goal", placing consent granularity on an equal footing with feature selection. Thirdly, course level learning analysis should not directly inherit the original logs on the platform side, but should be replaced by aggregated statistics or privacy protection analysis. The efficient use of AI in international communication English teaching can only be sustainable when the governance strength is sufficient to change the data path.

4 Conclusion

This thesis puts emphasis on the usage of AI and private protection in English international communication teaching, and builds a uniform proof framework for scenes, data objects, management behaviors. According to 82 researches and 112 scene records, this thesis places teaching effect, private information leak, management intensity and protection load into the same analysis dimension, hence proving that "high-efficiency use" and "strong controllability" in international communication English classrooms do not have natural overlapping, but this hence depends on whether the data route is carried out redesign.

(1) This article carries out decomposition on international email negotiations, public briefings, forum hosting, proposal dissemination, localization, and crisis simulation scenarios into comparable recording units on the organizational level, and binds and encodes AI modalities, teaching objectives, data objects, and governance measures, thus to avoid that single "language learning effect" covers up task differences.

(2) This article demonstrates in terms of methods and results that efficient applications mainly focus on text revision, audience adaptation, and localized rewriting, while oral expression and continuous interaction scenarios have higher exposure and heavier protection burdens; Meanwhile, only when local preprocessing, fine-grained authorization, and teacher auditing enter the implementation chain, can exposure growth be transformed into stable revenue.

(3) This article also has boundaries: the comprehensive evidence relies on the quality of publicly available research reports, and cannot fully recover the details of classroom governance that have not been made public; The compliance requirements of curriculum systems and platforms in different regions can also affect the location of the scene. Subsequent work needs to further validate the three-dimensional governance response surface based on real course data, and incorporate changes in student trust, cross platform deployment costs, and long-term learning outcomes into the same evaluation framework.

Funding

1.This work was supported by the 2024 General Project of Hunan Provincial Commission for Social Sciences Evaluation (Grant No. XSP24YBC233), titled “*Construction of a Quality Evaluation System for Deep Integration of Party Building and Institutional Development in Universities Based on Fuzzy Analytic Hierarchy Process*”, administrated by the Office of Hunan Provincial Commission for Social Sciences Evaluation, China.

2. This study was supported by the 2023 General Project of Teaching Reform in Hunan Provincial Higher Education Institutions (Grant No. HNJG-20231541), titled “*Research and Practice on Cultivating College Students' Foreign Language Digital storytelling Competence from the Perspective of Moral Education*”, funded by the Hunan Provincial Department of Education, China.

3. This research was supported by the 2024 General Project of Hunan Provincial Commission for Social Sciences Evaluation (Grant No. XSP24YBC246), titled “*Effects of Cognitive Efficiency on Rendering Quality in Chinese-English Translation*”, administered by the Office of Hunan Provincial Commission for Social Sciences Evaluation, China.

4. This study was supported by the 2024 General Project of Teaching Reform in Hunan

Provincial Higher Education Institutions (Grant No. 202401001852), titled "*Research and Practice on the Reform of College English Translation Teaching Assessment Based on Generative Artificial Intelligence*", funded by the Hunan Provincial Department of Education, China.

5. This work was supported by the 2025 Key Project for Graduate Research and Innovation of the Foreign Studies College, Hunan Normal University (Grant No. 2025-5), titled "*A Comparative Study on the Cognitive Mechanisms of Metaphor Translation in Political Discourse by Chinese and Foreign Large Language Models*".

About the Author

Haiyan Xiao:Associate Professor, School of General Education, Hunan University of Information Technology, Changsha, China; Key Researcher, Hunan Provincial Digital Intelligence Research Center for International Communication, Changsha, China

Tianyou Cen:PhD Student, Foreign Studies College, Hunan Normal University, Changsha, China; Lecturer, Hunan University of Information Technology, Changsha, China; Key Researcher, Hunan Provincial Digital Intelligence Research Center for International Communication, Changsha, China

Corresponding author. E-mail: centianyou@hunnu.edu.cn

References

[1] Ji, H., Han, I., & Ko, Y. (2023). A systematic review of conversational AI in language education: Focusing on the collaboration with human teachers. Journal of Research on Technology in Education, 55(1), 48–63.

[2] Liang, J.-C., Hwang, G.-J., Chen, M. R. A., et al. (2023). Roles and research foci of artificial intelligence in language education: An integrated bibliographic analysis and systematic review approach. Interactive Learning Environments, 31(7), 4270–4296.

[3] Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. Education and Information Technologies, 28(12), 15873–15892.

[4] UNESCO. (2023). Guidance for generative AI in education and research. UNESCO.

[5] UNESCO. (2023). Global education monitoring report 2023: Technology in education: A tool on whose terms? UNESCO.

[6] Lo, C. K., Yu, P. L. H., Xu, S., et al. (2024). Exploring the application of ChatGPT in ESL/EFL education and related research issues: A systematic review of empirical studies. Smart Learning Environments, 11(1), 50.

[7] Lai, W. Y. W., & Lee, J. S. (2024). A systematic review of conversational AI tools in ELT: Publication trends, tools, research methods, learning outcomes, and antecedents. Computers and Education: Artificial Intelligence, 7, 100291.

[8] Lee, S., Jeon, J., McKinley, J., et al. (2025). Generative AI and English language teaching: A global Englishes perspective. Annual Review of Applied Linguistics, 45, 85–108.

[9] Liu, Y., Zhang, H., Jiang, M., et al. (2024). A systematic review of research on emotional artificial intelligence in English language education. System, 126, 103478.

[10] Han, J., & Li, M. (2024). Exploring ChatGPT-supported teacher feedback in the EFL context. System, 126, 103502.

[11] Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. System, 124, 103362.

[12] Duong, T. N. A., & Chen, H. L. (2025). An AI chatbot for EFL writing: Students' usage tendencies, writing performance, and perceptions. Journal of Educational Computing Research, 63(2), 406–430.

[13] Moorhouse, B. L., Wan, Y., Wu, C., et al. (2024). Developing language teachers' professional generative AI competence: An intervention study in an initial language teacher education course. System, 125, 103399.

[14] Abdelhalim, S. M., Alsahil, A. A., Baek, C., et al. (2025). University EFL instructors' engagement with generative AI: A cross-cultural analysis. System, 135, 103861.

[15] Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. International Journal of Educational Technology in Higher Education, 20(1), 38.

[16] Liu, Q., & Khalil, M. (2023). Understanding privacy and data protection issues in learning analytics using a systematic review. British Journal of Educational Technology, 54(6), 1715–1747.

[17] Mutimukwe, C., Viberg, O., Öberg, L. M., et al. (2022). Students' privacy concerns in learning analytics: Model development. British Journal of Educational Technology, 53(4), 932–951.

[18] Viberg, O., Kizilcec, R. F., Jivet, I., et al. (2024). Cultural differences in students' privacy concerns in learning analytics across Germany, South Korea, Spain, Sweden, and the United States. Computers in Human Behavior Reports, 14, 100416.

[19] Liu, Q., Shakya, R., Khalil, M., et al. (2025). Advancing privacy in learning analytics using differential privacy. In Proceedings of the 15th International Learning Analytics and Knowledge Conference (pp. 181–191).

[20] Huang, Z., & Cao, C. (2025). Federated learning for AI-assisted education: Privacy-preserving and cross-platform collaborative modeling in higher education. In Proceedings of the 2nd International Conference on Machine Intelligence and Digital Applications (pp. 146–151).